

EVALUATING THE IMPACT OF A MULTISENSORY PHONICS CURRICULUM ON PHONEMIC AWARENESS IN CHINESE KINDERGARTEN EFL LEARNERS: A MIXED-METHODS STUDY

Marisha Sookdeo
marishasookdeo24@gmail.com

Abstract

A quasi-experimental study was conducted to evaluate the effectiveness of a multisensory phonics curriculum in enhancing phonemic awareness among Chinese kindergarten learners of English as a Foreign Language (EFL), in comparison to a standard language-integrated approach. Conducted over 4 months, the study involved 30 students from an international, bilingual school in Beijing, China, divided evenly into an experimental group (phonics intervention) and a control group (no intervention). Using an explanatory sequential mixed-methods design framework, the study combined quantitative data from test scores and ratings with qualitative insights from class observations and parental feedback. Mann-Whitney U-test results indicated improvement in phonemic awareness by increased post-test scores in alphabet letter-sound recognition, initial letter sound proficiency, and vocabulary acquisition. The results demonstrated that the multisensory phonics curriculum significantly enhanced phonemic awareness in the experimental group.

Keywords: multisensory phonics, phonemic awareness, English as a Foreign Language, kindergarten education, Chinese learners

Background of the Study

In today's globalized world, English as a foreign language (EFL) is crucial, especially in China, where English proficiency is a gateway to academic, career, and cultural opportunities (Liu, 2022). Early education, particularly in kindergarten, is vital for developing linguistic skills. Consequently, phonics programs have become popular in Chinese kindergartens due to the emphasis on early English proficiency (Chen et al., 2022; Huo & Wang, 2017).

Phonics instruction, which links sounds with letters, is fundamental for learning to read and write, particularly for young learners (Shin & Crandall, 2018). The National Reading Panel (2000) asserts that understanding sound-symbol correspondence equips children to decode unfamiliar words, fostering independent reading and language proficiency. In China's multilingual context, phonics instruction provides a structured approach to language acquisition.

Despite its importance, effective phonics instruction in Chinese kindergartens faces challenges. Most existing phonics programs, such as Oxford Phonics and Scholastic Phonics, are

designed for native English speakers and may not suit Chinese students' needs, hindering effective phonics acquisition (Huo & Wang, 2017; Sun & Xie, 2021). Research is needed to identify methods that align with Chinese linguistic patterns and cultural learning styles, optimizing language learning for young learners (Chen et al., 2022; Jiang et al., 2023; Sun et al., 2016). Multisensory phonics, engaging multiple senses simultaneously, could fill this gap.

Multisensory approaches in phonics instruction involve visual, auditory, and kinesthetics elements, potentially enhancing learning outcomes for young learners (Attia, 2020; Shin & Crandall, 2018). Research indicates that these approaches, tailored for Early Childhood Education (ECE), create dynamic and inclusive learning environments, making phonics acquisition more effective and enjoyable (Attia, 2020; Bdeir et al., 2022; Chen et al., 2022; Huo & Wang, 2017). However, there is a lack of research on multisensory phonics instruction in Chinese kindergarten settings, necessitating an investigation into its impact on early literacy development (Chen et al., 2022; Liu, 2022).

Statement of the Problem

Despite the integration of a 5-minute phonics session in English lessons, its effectiveness in achieving desired learning outcomes is uncertain. Variations in teacher expertise and the brief nature of these sessions raise questions about their adequacy. The current method lacks a definitive structure and fails to address individual learning styles, limiting engagement and retention of phonetic concepts. It prioritizes decoding over meaning-making, which is essential for engaging deeply with language (Jiang et al., 2023). Effective phonics instruction should emphasize letter recognition, corresponding sounds, initial sound proficiency, and overall vocabulary acquisition (Chen et al., 2022; Huo & Wang, 2017).

The current 5-minute method's limitations underscore the need for a more differentiated, multisensory approach that caters to diverse learning styles and provides extended exposure to phonics instruction. Therefore, this study was conducted to investigate whether a multisensory phonics approach can better facilitate phonemic awareness and language acquisition in Chinese EFL kindergarten students.

Significance of the Study

This study addresses current limitations and research gaps in phonics instruction for Chinese kindergarten EFL learners. By examining the deficiencies of the 5-minute method and the potential benefits of multisensory instruction, the study aims to bridge the gap between existing phonics instruction and the needs of young learners. The findings could inform the development of more effective, culturally relevant phonics programs.

The findings of this study could be beneficial to various educational stakeholders. Teachers might gain knowledge of more effective phonics methods, and policymakers could use the findings to reform curricula, making them more attuned to local linguistic and cultural nuances.

Research Questions

The purpose of this study was to evaluate the effectiveness of a multisensory phonics curriculum in enhancing phonemic awareness among Chinese EFL kindergarten students, compared to a 5-minute language-integrated approach. The curriculum involves 20-minute daily lessons over 12 weeks, focusing on teaching four to five letters and sounds with corresponding vocabulary each month.

The overall objective of the study was to examine several specific aspects of the phonics intervention's effectiveness; thus, the corresponding research questions were as follows.

1. What is the impact of a structured 20-minute daily phonics lesson on the ability of 4 to 5-year-old Chinese EFL students to recognize alphabet letters and their corresponding sounds?
2. How does a structured multisensory phonics curriculum affect the students' proficiency in sounding out initial letter sounds in words compared to a language-integrated approach?
3. To what extent does a targeted phonics curriculum enhance the overall vocabulary acquisition of Chinese EFL students aged 4 to 5?

Scope of the Study

The study was conducted at an international, bilingual Montessori kindergarten in Beijing, China, involving 30 Chinese EFL students aged 4 to 5 years old. The sample represented a homogeneous group in terms of linguistic background, age, educational setting, and socioeconomic status, enhancing the study's internal validity (Creswell & Guetterman, 2018).

The study took 4 months to complete and both the control and experimental groups were taught by the same teacher to minimize instructional variability and reduce the risk of confounding variables.

Literature Review

Phonemic awareness, the ability to recognize and manipulate sounds in language at the phoneme level, is a foundational skill in language acquisition. Essential for reading, spelling, and language comprehension, the lack of phonemic awareness, however, can impede decoding words and understanding text. EFL learners in Chinese kindergartens face unique challenges in developing phonemic awareness due to linguistic and cultural differences. This literature review explores the effectiveness of multisensory phonics in addressing these deficits among Chinese kindergarten EFL learners, identifying gaps in existing research, and establishing a framework for further study.

The Concept of Multisensory Phonics

Learning a second language at an early age enhances cognitive flexibility, problem-solving skills, and cultural awareness (Johnson, 2019; NRP, 2002; Paradis, 2011). Traditional teaching methods, however, often fail to engage young EFL learners effectively. They typically rely on visual aids and rote learning, neglecting auditory and kinesthetics styles, and lack interactivity, which can disengage children. Multisensory learning, which engages auditory, visual, and kinesthetics senses simultaneously, offers a promising alternative. This approach makes learning more accessible and enjoyable by incorporating songs, interactive games, role-playing, and hands-on materials, thus fostering deeper understanding and retention of language (Attia, 2020).

Previous Research on Phonics Instruction

Phonics instruction has been extensively studied globally, revealing varied approaches based on educational contexts and cultural factors. Studies in Asia indicate the efficacy of phonics-based methods in enhancing early literacy skills. Key areas of phonics instruction include:

- **Alphabet Letter-Sound Recognition:** Studies by Bdeir et al. (2022) suggest that incorporating familiar sounds from learners' native languages and using explicit rime-based methods can enhance alphabet recognition. Consistent, systematic phonics teaching over one or two years positively impacts early literacy outcomes.
- **Initial Letter Sound Proficiency:** Tailored phonics instruction that addresses individual needs improves initial letter sound proficiency. Explicit instruction and phonemic awareness activities such as those proposed by Agegnehu et al. (2023), are crucial in helping learners associate letters with sounds.
- **Vocabulary Acquisition:** Phonics instruction interlinked with phonological awareness significantly contributes to vocabulary development (Yeung et al., 2013). Integrating phonics with whole language approaches provides a broader context for word recognition and understanding.

Gaps in Existing Literature

Despite the insights gained, there are notable gaps in research, particularly in China:

- **Age-Specific Research:** There's a dearth of comprehensive studies focusing on kindergarten-aged EFL learners in China, with most research encompassing broader age ranges (Chen et al., 2022).
- **Linguistic and Cultural Context:** Understanding the impact of linguistic and cultural differences on phonics instruction is limited (Jiang et al., 2023). Research tailored to Chinese learners' unique phonological systems is needed.
- **Long-Term Studies:** Most studies are short-term. Longitudinal research tracking learners' progress over several years is essential for understanding the sustained impact of phonics instruction (Shenoy et al., 2022).
- **Diverse Learner Profiles:** More research is required on how different phonics approaches cater to diverse learners, including those with learning disabilities (Liu, 2022).

- Standardized Implementation and Teacher Training: Effective implementation of phonics programs and adequate teacher training are critical areas needing more exploration.
- Methodological Rigor: Existing studies often have methodological flaws. Robust research designs are necessary to establish stronger causal relationships between phonics instruction and its effectiveness (Huo & Wang, 2017).

Theoretical Framework

The theoretical framework for this study integrates several educational theories:

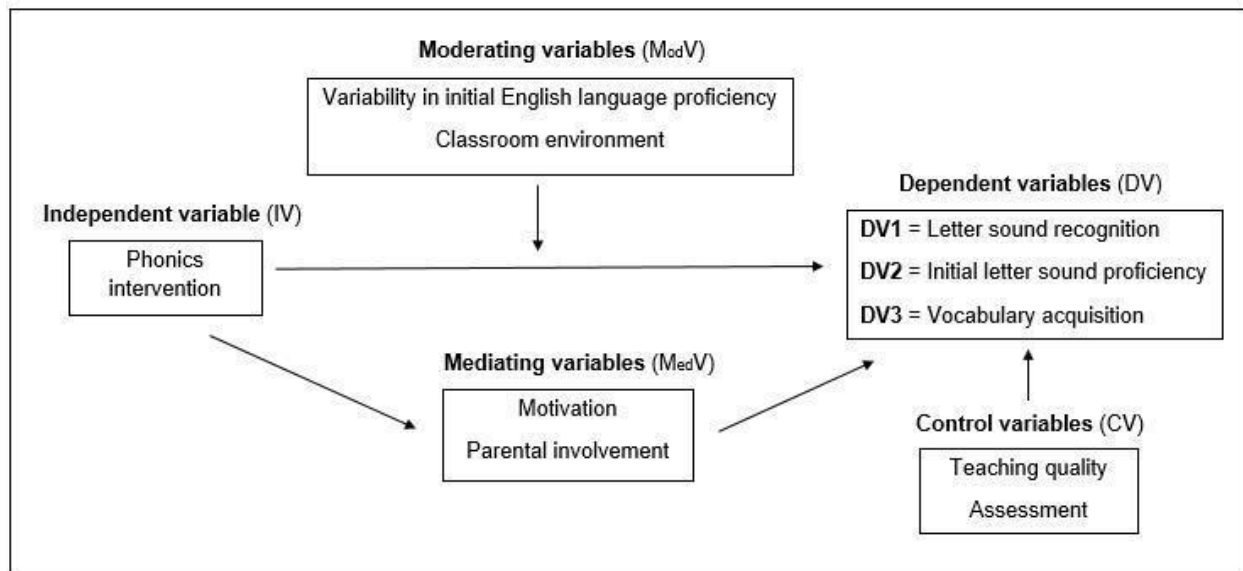
- Bottom-Up Theory of Reading Process: Emphasizes starting from basic visual information and systematically progressing through linguistic units (Chen et al., 2022).
- Behaviourist Learning Theory: Supports repetitive practice and systematic phonics instruction to reinforce letter-sound associations (Johnson, 2019).
- Constructivist Learning Theory: Advocates for learner-centred instruction, where learners discover decoding rules and integrate new knowledge (Pritchard, 2017).
- Piaget's Cognitive Development Theory: Focuses on teaching phonics according to learners' cognitive development stages (Sun et al., 2016).
- Zone of Proximal Development: Stresses guided learning, where teachers assist learners in tasks beyond their independent capabilities (Johnson, 2019).
- Comparative Effectiveness Theory: Promotes empirical comparison of different instructional methods to identify the most effective approaches (Mertler, 2021).

Conceptual Framework

The conceptual framework for this study outlines the relationships among key variables in investigating the effects of phonics intervention on phonemic awareness in Chinese EFL kindergarten settings (see Figure 1). The independent variable is the phonics intervention, and the dependent variables are letter sound recognition, initial letter sound proficiency, and vocabulary acquisition. Variances in initial English proficiency and classroom environment act as moderating variables, influencing the relationship dynamics.

Figure 1

Conceptual Framework for the Study (Patten, 2017)



In summary, while phonological-based instruction shows promise in enhancing foundational literacy skills among EFL learners in China, further research is necessary to address methodological shortcomings and provide more definitive evidence of its effectiveness.

Research Methodology

The study utilized a quasi-experimental pretest and post-test designed to compare two groups within a kindergarten setting where random assignment was not feasible. This design involved an experimental group that received an intervention and a control group that did not, both of which were pretested and post tested. This method allowed for a clear demonstration of cause-and-effect relationships between the intervention and improvements in phonemic awareness, suitable for real-world classroom settings with fixed structures (Cohen et al., 2017).

- Location and Sample of the Study

The study adopted a non-probability sampling technique combining purposive and convenience sampling. It targets 30 Chinese EFL students aged 4-5 years, ensuring the sample size is manageable and statistically significant. The selection of two classes, each with 15 students, is based on criteria like Montessori-centric curriculum and similar teaching environments, ensuring homogeneity and minimizing sampling bias.

The school's curriculum was meticulously designed to integrate the principles of Montessori education, Chinese educational policy, and the CEFR (Common European Framework of Reference for Languages) international standard for language learning. English instruction forms a significant component of the curriculum, with students receiving approximately three hours of English lessons daily. This includes a 30-minute whole-class session in the morning, followed by 15-minute group lessons for smaller cohorts of 4 to 5 students. Additionally, individualised Montessori and Chinese lessons are seamlessly integrated into the morning teaching procession, providing students with a well-rounded educational experience.

Despite the emphasis on English language acquisition, Chinese remains the primary medium of communication among teachers and between students and the Chinese teaching team outside of English lesson times. This reflects the school's commitment to preserving students' native language and cultural heritage while fostering proficiency in English as a global language of communication and commerce.

In terms of demographics, all intended participants for this study, totalling 30 Chinese EFL students aged 4 to 5 years old, were successfully enrolled and utilised in the research. The selection process adhered to strict criteria to ensure a sufficient representative sample of the target population (CI = 95%, $\alpha = 0.05$, E = 1 %). Each participant was chosen based on their eligibility, including age, developmental needs, linguistic background (native Chinese speakers), English language acquisition proficiency level, and parental consent. The participants were drawn from the kindergarten's student body, reflecting the diverse socioeconomic backgrounds and familial support structures characteristic of the institution.

The sample was evenly divided between the control group and an experimental group, each comprising 15 participants representing 50% of the total sample population. In the control group, 53.33% were male (8 individuals) and 46.67% were female (7 individuals). Likewise, in the experimental group, 53.33% were male (8 individuals) and 46.67% were female (7 individuals). This indicates that both the control and experimental groups were balanced in terms of gender representation, which is important for ensuring that the results are not biased towards one gender. This balance further enhances the validity of the study's findings by minimising the potential influence of gender-related factors on the results.

- Phonics Curriculum Intervention Plan

The intervention involves a multisensory phonics curriculum that introduces 4-5 letters and their sounds each month, using a blend of auditory, visual, and kinesthetics learning techniques. It incorporates phonics songs, rhyming games, and story listening for auditory recognition, while employing images, videos, gestures, and interactive whiteboards for visual reinforcement. Physical activities such as letter tracing and word puzzles engage tactile senses. Additionally, the program offers differentiated instruction to tailor learning experiences to individual needs and pacing.

- Instrumentation

The study employed adapted assessment tools to screen phonemic awareness, ensuring cultural and linguistic appropriateness. These included researcher-developed assessments, class observations, and parental feedback, tailored to measure letter-sound recognition, initial letter sound proficiency, and vocabulary acquisition.

- Validity and Reliability

Measures to ensure validity included a homogeneous sample population and consistent teaching quality. Reliability was assessed through Cronbach Alpha coefficients and inter-reliability checks (Cohen et al., 2017). The study also employed triangulation of data from various sources to corroborate findings and reduce biases (Cope & Kalantzis, 2015; Lund, 2012; Onwuegbuzie & Leech, 2004).

Data Collection Methods

An explanatory sequential mixed methods design was employed, prioritizing quantitative data and using qualitative data to enrich interpretations (Caracelli & Greene, 1993; Onwuegbuzie & Leech, 2004). Participants were anonymized with unique identifiers, and data included demographics, test scores, class observations, and parental feedback.

Quantitative Data

Collected quantitative data includes:

- Sample Demographics: Gender, initial English proficiency (measured by CEFR), exposure to English outside school, and motivation levels.
- Test Scores: Pretest and post-test scores on alphabet recognition, initial letter sound proficiency, and vocabulary acquisition.
- Class Observations: Ratings on eight criteria, recorded during phonics lessons.
- Parental Feedback: Structured ratings on seven questions related to phonemic awareness and student confidence in English.

Qualitative Data

Qualitative data from class observations and parental feedback was transcribed and analysed using thematic analysis techniques to identify recurring themes and patterns.

Data Analysis Process

The analysis integrated quantitative and qualitative data to evaluate the intervention's effectiveness. SPSS was used for statistical analyses, including descriptive statistics, normality tests, and t-tests or Mann-Whitney U tests based on data distribution. Qualitative data was analysed for themes related to student engagement and parental observations.

Ethical Considerations

The study adhered to ethical standards by prioritizing honesty, transparency, and integrity as advocated by the British Educational Research Association (BERA) (2018). It obtained necessary permissions and informed consent from schools and parents, emphasizing voluntary participation and the right to withdraw. To minimize stress, it ensured a supportive environment with age-appropriate tasks and maintained open communication with parents. Confidentiality was upheld by securing participants' data and anonymity through careful handling and the use of numeric identifiers.

Mitigating Researcher Bias

To minimize researcher bias, strategies include employing standardized protocols, blinding techniques, and data triangulation to ensure objectivity (Mertler, 2021; Patten, 2017).. Involving participants and parents in the research process helps empower them and provides diverse perspectives. Transparent reporting involves openly disclosing biases and documenting methodological choices.

Seeking peer feedback and collaboration from colleagues and supervisors is crucial, as is engaging in reflective practice to continually assess and adjust one's role and interactions to maintain objective data interpretation.

By employing these rigorous methodologies and ethical practices, the study aimed to provide reliable and valid insights into the effectiveness of the multisensory phonics curriculum for improving phonemic awareness in Chinese EFL kindergarten students

Research Findings

(i) Base Variables Related to Initial English Language Proficiency Levels

By comparing the distribution of initial English language acquisition proficiency levels between the control and experimental groups, the researcher can determine if there are any significant differences or similarities in the participants' language abilities. This comparison of language proficiency levels within and between groups helps ensure that the groups are reasonably comparable before the intervention is applied.

A frequency analysis of initial English Language acquisition proficiency levels for both control and experimental groups revealed that the initial English language proficiency for both the control and experimental groups predominantly fell within the elementary level (93.33%), each consisting of 14 individuals; with a small proportion of participants (1 individual for each group) at the pre-intermediate level (6.67%). No participants were at the beginner, intermediate, or upper intermediate levels. These findings indicate that there was minimal variability in proficiency levels between the two groups, as reflected by the identical distributions. This equivalence is important for ensuring that any differences observed in later assessments can be attributed to the intervention rather than initial proficiency disparities (Patten, 2017).

Frequency analysis also revealed that the majority of participants (90%) have limited exposure to English outside of school. Among them, 43.33% belong to the control group, while 46.67% are from the experiment group. Only a small proportion of participants (3.33%) receive English tuition rarely, with this category represented by one participant from the experiment group. Additionally, a very small percentage (6.67%) of participants receive English tuition occasionally, with two participants falling into this category, both from the control group. Notably, no participants reported receiving English tuition frequently.

The frequency of parental involvement can exert a significant influence on the results of the intervention by affecting baseline comparability, intervention implementation, generalisation of results, sustainability of gains, and mediating effects (Patten, 2017). To account for the potential effects of such, a frequency analysis on parental involvement during the intervention was

conducted. The findings indicate that within the control group, 6.67% of participants (2 individuals) received parental support "Once a week", whereas the remaining 43.33% (13 individuals) "Never" received any form of parental assistance.

Similarly, in the experimental group, 3.33% (1 individual) was categorised as having "Rarely" received parental support, with the majority, accounting for 46.67% (14 individuals), "Never" having received any parental support related to the study.

Understanding participants' motivation levels before the intervention provides context for understanding any changes observed post-intervention and helps distinguish between the effects of the intervention and other factors influencing motivation. This allows the researcher to assess the effectiveness of the intervention by comparing post-intervention motivation levels to those of pre-intervention. The results indicate that prior to the intervention, students in the control group exhibited predominantly low motivation levels, with 100% reporting feeling not "not motivated". In contrast, in the experimental group, while the majority (86.67%) also reported being "not motivated", a small proportion (13.33%) indicated feeling "somewhat motivated". Notably, none of the students in either group reported feeling "very motivated" before the intervention. Overall, among the total sample of 30 students, 93.33% expressed low motivation levels, while only 6.67% reported feeling somewhat motivated, with no students indicating a high level of motivation. This suggests a significant need for intervention strategies aimed at enhancing student motivation across both groups (Johnson, 2019; Paradis, 2011; Sun et al., 2016).

After the intervention, results indicate notable disparities between the control and experimental groups. In the control group, 40% of students (12 individuals) reported being "not motivated," while only 6.67% of students (2 individuals) in the experimental group expressed similar sentiments. Conversely, the experimental group displayed substantially higher levels of motivation, with 33.33% (10 individuals) indicating they were "very motivated" compared to none in the control group. Additionally, while 10% of students (2 individuals) in both groups reported feeling "somewhat motivated," the experimental group had a higher percentage of students falling into this category. This suggests that the intervention not only mitigated feelings of low motivation but also fostered a greater sense of enthusiasm and engagement among participants, particularly evident in the experimental group (Pritchard, 2017).

A Shapiro-Wilk test was performed and showed that the distribution of student motivation before intervention for the experimental group significantly deviated from normality, $W(15) = .41$, $p < .001$. Similarly, a Shapiro-Wilk test showed that the distribution of student motivation after intervention for the experimental group significantly deviated from normality, $W(15) = .66$, $p < .001$. Based on these outcomes, Mann-Whitney U-Tests were then conducted (see Table 1).

The results of the Mann-Whitney U-test indicated that student motivation before intervention for the experimental group (mean rank = 16.50) was 12.12% slightly higher than that of the control group (mean rank = 14.50), $U = 97.5$, $z = 1.44$, $p = .539$, *ns*, with a small effect size $r = 0.26$. The test indicated no statistically significant difference in student motivation before the intervention between the control and experimental groups, as the p-value was well above the 0.05 significance threshold.

Table 1*Mann-Whitney U-test Results on Motivation Levels*

Variables	Control		Experiment		U	z	p
	n	Mean rank	n	Mean rank			
Student motivation before intervention	15	14.50	15	16.50	95.5	-1.44	0.539
Student motivation after intervention	15	9.50	15	21.50	22.5	-4.04	< .001

The Mann-Whitney U-test results indicated that student motivation after intervention for the experimental group ($Mdn = 2.53$, mean rank = 21.50) was 55.81% higher than that of the control group ($Mdn = 1.20$, mean rank = 9.50), $U = 22.5$, $z = 4.04$, $p < .001$, with a large effect size, $r = 0.74$. The test revealed a statistically significant difference in student motivation after the intervention, with the experimental group showing significantly higher motivation than the control group. The p-value was far below the 0.05 significance threshold. This suggests that the intervention had a substantial positive impact on the motivation of participants in the experimental group compared to those in the control group.

(ii) Differences in Test Scores

All sample participants' ($N = 30$) pretest and post-test scores for each dependent variable: letter recognition (LR) and letter-sound recognition (DV1), initial letter sound proficiency (DV2), and vocabulary acquisition (DV3), were measured. Pretest and post-test scores were collected from both the experimental group (multisensory phonics curriculum) and the control group (language-integrated phonics curriculum).

The difference in pretest and post-test scores for each respective dependent variable significantly increased across the experimental group than when compared to the control group, as depicted in Figure 2. This graphical illustration (violin plot) suggests that the multisensory phonics curriculum had a substantial effect on the participants' average test scores across all dependent variables, indicating its potential significance in influencing the measured outcome of improving phonemic awareness.

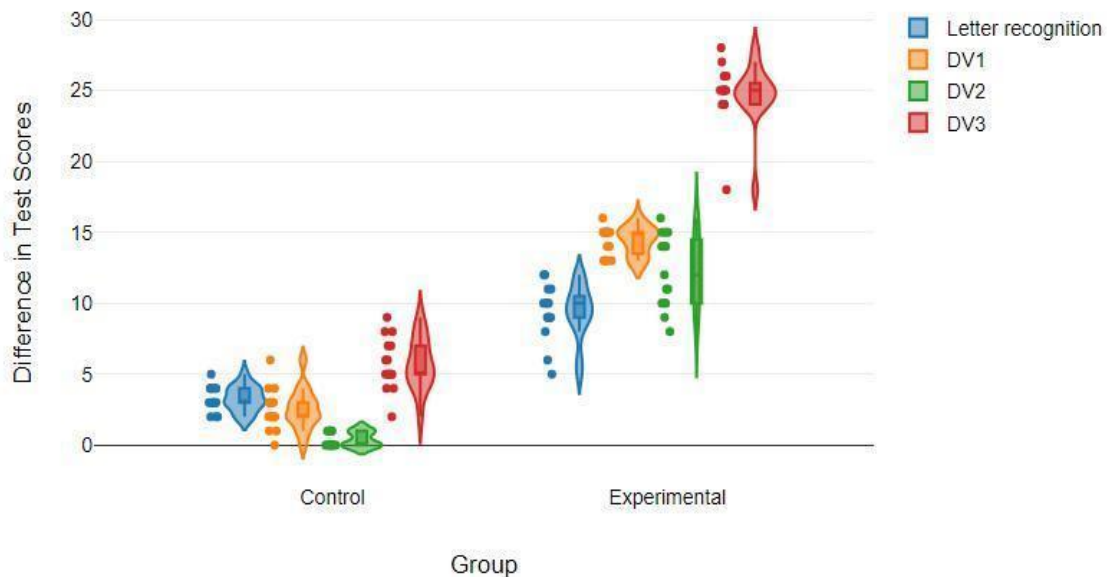
Cronbach's alpha for the twelve items (pretests, post-tests, and their differences) measuring the effectiveness of the multisensory phonics curriculum in improving phonemic awareness was .93, suggesting an excellent level of internal consistency. The average test score, denoted by TS.Data_Effect_MPC, was calculated by dividing the difference in test scores (pretest and post-test scores) for each dependent variable by the maximum value then totalling them.

The experimental group's average test scores ($M = 3.03$, $SD = 0.27$) indicated a significant increase in phonemic awareness when compared to the control group's average test scores ($M = 0.57$, $SD = 0.14$). A Mann-Whitney U test was conducted to compare the average test scores

(TS.Data_Effect_MPC) on the effectiveness of the multisensory curriculum in improving phonemic awareness between the experimental and control groups (see Table 4.2).

Figure 2

Collective Difference in Test Scores by Group Comparison



The results indicated that the average test scores for participants in the experimental group ($Mdn = 3.09$, mean rank = 23.00) were 81.89% higher than that of participants in the control group ($Mdn = 0.56$, mean rank = 8.00), $z = 4.67$, $p < .001$, with a large effect size $r = 0.85$. These results provide strong evidence that the multisensory curriculum effectively enhances phonemic awareness among participants, demonstrating a substantial and statistically significant improvement in average test scores for the experimental group.

Table 2

Mann-Whitney U-Test Results on Average Test Scores

Variable	Control		Experiment		r	z	p
	n	Mean rank	n	Mean rank			
Test scores	15	8.0	15	23.0	0.85	-4.67	<.001

(iii) Results and Findings from Class Observations

The reliability of the eight criteria-rated items in measuring the class observation rating scores on the effectiveness of the multisensory phonics curriculum in improving phonemic awareness was found to have a Cronbach's alpha of .98. This high coefficient indicates an excellent level of internal consistency. The average class observation ratings, denoted by CO.Data_Effect_MPC, was calculated by dividing the rating for each criterion by the total expected value of 5 then totalling them.

Observational data, quantified through a Likert scale rating criteria (1 = low, 5 = high), showed higher average class observation ratings (59.14%) for the experimental group ($M = 4.65$, $SD = 0.26$) compared to the control group ($M = 1.90$, $SD = 0.48$). Owing to a small sample size ($N = 10$), determining the distribution of the average class observation ratings was important for choosing an appropriate statistical method. A Shapiro-Wilk test was performed and showed that the distribution of average class observation ratings significantly deviated from normality, $W(30) = .82$, $p = .021$. Based on this outcome, a Mann-Whitney U test was then conducted to compare the average class observation ratings on the effectiveness of the multisensory in improving phonemic awareness between the experimental and control groups (see Table 3).

The results indicated that average class observation ratings for the experimental group ($Mdn = 4.63$, mean rank = 8.00) were 54.90% higher than that of the control group ($Mdn = 1.88$, mean rank = 3.00), $z = 2.63$, $p = .008$, with a large effect size $r = 0.83$. These results suggest that the multisensory approach had a substantial impact on improving phonemic awareness in the experimental group compared to the control group.

Table 3

Mann-Whitney U-Test Results on Class Observation Rating

Variable	Control		Experiment		<i>r</i>	<i>z</i>	<i>p</i>
	<i>n</i>	Mean rank	<i>n</i>	Mean rank			
Ratings	15	3.0	15	8.0	0.83	-2.63	.008

Qualitative findings from the class observations were derived from ten observation sheets, split evenly between control and experimental groups. A thematic analysis identified recurring differences in instructional strategies and observations between the class groups, as summarised in Table 4. The results revealed that the experiment group, using multisensory strategies, small group instruction, collaborative learning, interactive games, strong differentiation and extensive use of technology, showed significantly better student engagement, attention span and participation, compared to the control group, which used traditional methods, whole class instruction, direct instruction, drill practice activities, limited differentiation and minimal technology.

Table 4*Observed Thematic Differences between Class Groups*

Theme	Control	Experiment
Instructional strategies	Traditional method	Multisensory method
	Whole class instruction	Small group instruction
	Direct instruction	Collaborative learning
	Drill practice activities	Interactive games
	Limited differentiation	Strong differentiation
	Minimal use of technology	Extensive use of technology
Observations	Low student engagement	High student engagement
	Short attention span	Long attention span
	Limited participation	Active participation

(iv) Findings from Parental Feedback

The reliability of seven items in measuring parental feedback on the effectiveness of the multisensory phonics in improving phonemic awareness, was assessed. The internal consistency of these items was found to be at a good level, with a Cronbach's alpha of .84. The average parental feedback rating scores, denoted by PF.Data_Effect_MPC, was calculated by dividing the rating for each question by their maximum value and then totalling them.

The mean rating scores for the experimental group ($M = 16.88$, $SD = 0.96$) indicated a significant increase in phonemic awareness when compared to that of the control group ($M = 4.37$, $SD = 0.8$). A Shapiro-Wilk test was then performed and showed that the distribution of average rating scores departed significantly from normality, $W(30) = .73$, $p < 0.01$). A Mann-Whitney U test was conducted to compare the average rating scores from parental feedback on the effectiveness of the multisensory phonics curriculum in improving phonemic awareness between the experimental and control groups (see Table 5).

The results indicated the average rating scores for the experimental group ($Mdn = 17.00$, mean rank = 23.00) were 73.53% higher than those of the control group ($Mdn = 4.50$, mean rank = 8.00), $z = 4.68$, $p < .001$, with a large effect size, $r = 0.85$. This suggests that parents perceived the multisensory phonics curriculum as highly effective in improving phonemic awareness.

Table 5*Mann-Whitney U-Test Results on Parental Feedback Ratings*

Variable	Control		Experiment		<i>r</i>	<i>z</i>	<i>p</i>
	<i>n</i>	Mean rank	<i>n</i>	Mean rank			
Ratings	15	8.0	15	23.0	0.85	-4.68	<.001

Qualitative findings from parental feedback were derived from eighteen written responses on the parent feedback questionnaire ($N = 18$), for both class groups. Including both groups enhances the validity of the findings. It ensures that the results are not biased or limited to a specific group and provides a more comprehensive understanding of parental feedback (Cohen et al., 2017; Patten, 2017). A thematic analysis captured a wide range of parental experiences and viewpoints, providing a more holistic understanding of the suggestions and concerns raised (see Table 6).

The results indicated that a significant portion of parents (72.22%) want more interactive phonics activities, indicating a strong preference for engaging and hands-on learning methods. Half of the parents (50.00%) desire additional phonics practice resources, suggesting a need for more materials to support their children's learning at home. A notable number (38.89%) see value in parent workshops, which could help them support their children's phonics development more effectively. Over half (55.56%) are interested in receiving progress reports on their children's phonics development, highlighting a need for regular updates and communication.

A smaller group (16.67%) were concerned about the screen time associated with online resources, reflecting apprehensions about digital learning tools. Some parents (27.78%) believe there should be more individualised support for struggling learners, indicating a demand for personalised attention to address diverse learning needs.

Table 6*Thematic Suggestions and Concerns from Parent Feedback*

Theme	Description	<i>n</i>	%
Suggestions	Provide more interactive phonics activities	13	72.22
	Provide more phonics practice resources	9	50.00
	Offer parent workshops on effective phonics	7	38.89
	Progress reports on phonics development	10	55.56
Concerns	Screen time associated with online resources	3	16.67
	More individualised support for struggling learners	5	27.78

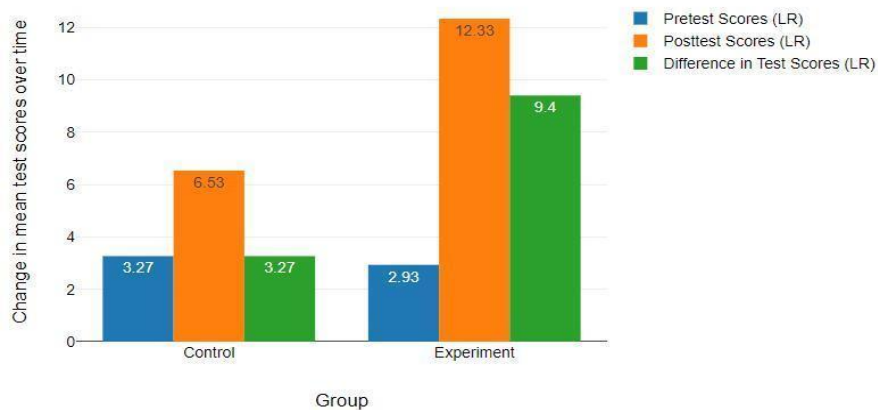
(iv) Effect on letter recognition (LR)

Based on the analysis of the pretests scores, post-test scores and their differences, it was found that letter recognition (LR) was significantly higher in the experimental group than the control group, as graphically indicated in Figure 3. The pretest scores for both groups were fairly similar: the control group ($n = 15$) had a mean of 3.27 ($SD = 1.58$, 95% $CI = 2.39, 4.14$), while the experimental group ($n = 15$) had a mean of 2.93 ($SD = 0.96$, 95% $CI = 2.4, 3.47$). The post-test scores showed an increase for both groups: the control group had a mean of 6.53 ($SD = 1.46$, 95% $CI = 5.73, 7.34$), and the experimental group had a significantly higher mean of 12.33 ($SD = 1.88$, 95% $CI = 11.29, 13.37$). The difference in test scores, representing the increase from pretest to post-test, for the control group had a mean of 3.27 ($SD = 0.88$, 95% $CI = 2.78, 3.76$), while the experimental group showed a substantially greater increase with a mean of 9.4 ($SD = 1.96$, 95% $CI = 8.32, 10.48$).

A Shapiro-Wilk test was then performed and showed that the distribution of test scores for letter recognition departed significantly from normality, $W(30) = .87$, $p = .002$. A Mann-Whitney U test was conducted to compare letter recognition test scores between the experimental and control groups (see Table 7).

Figure 3

Group Comparison of Mean Test Scores for Letter Recognition



The results indicated letter recognition test scores (LR) for the experimental group ($Mdn = 10.00$, mean rank = 22.97) were 70% higher than those of the control group ($Mdn = 3.00$, mean rank = 8.03), $z = 4.69$, $p < .001$, with a large effect size, $r = 0.86$. This suggests that the multisensory phonics curriculum was highly effective. The results were robust enough to be statistically significant even when accounting for the non-normal distribution of the data (Cohen et al., 2017).

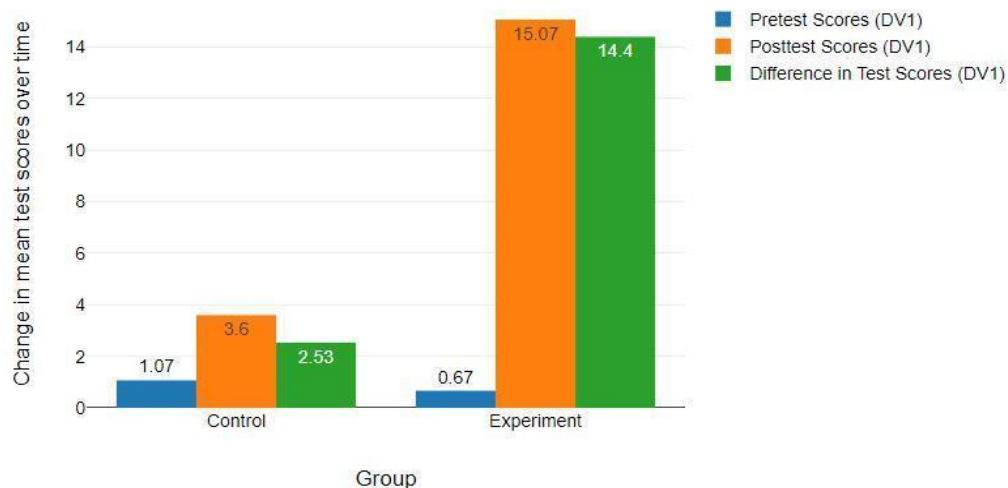
Table 7
Mann-Whitney U-Test Results on LR Test Scores

Variable	Control		Experiment		<i>r</i>	<i>z</i>	<i>p</i>
	<i>n</i>	Mean rank	<i>n</i>	Mean rank			
(LR)	15	8.03	15	22.97	0.86	-4.69	<.001

(v) Effect on letter-sound recognition (DV1)

Based on the pretests scores, post-test scores and their differences for the dependent variable, it was found that letter-sound recognition (DV1) was significantly higher in the experimental group than the control group, as graphically indicated in Figure 4. The pretest scores for the control group ($M = 1.07$, $SD = 0.88$, 95% $CI = 0.58, 1.56$) and the experimental group ($M = 0.67$, $SD = 0.62$, 95% $CI = 0.32, 1.01$) indicated that there was not much difference between the two groups before the intervention. Post-test scores indicated a significant increase in the experimental group ($M = 15.07$, $SD = 0.70$, 95% $CI = 14.68, 15.46$) compared to a smaller increase in the control group ($M = 3.60$, $SD = 1.24$, 95% $CI = 2.91, 4.29$). The difference in test scores further highlighted the improvement in the experimental group ($M = 14.40$, $SD = 0.99$, 95% $CI = 13.85, 14.95$) over the control group ($M = 2.53$, $SD = 1.46$, 95% $CI = 1.73, 3.34$).

Figure 4
Group Comparison of Mean Test Scores for Letter-sound Recognition



A Shapiro-Wilk test was then performed and showed that the distribution of test scores for letter-sound recognition departed significantly from normality, $W(30) = .79, p < .001$. A Mann-Whitney U test was conducted to compare letter-sound recognition test scores between the experimental and control groups (see Table 8). The results indicated letter-sound recognition test scores (DV2) for the experimental group ($Mdn = 15.00$, mean rank = 23.00) were 86.67% higher than those of the control group ($Mdn = 2.00$, mean rank = 8.00), $z = 4.73, p < .001$, with a large effect size, $r = 0.86$.

Table 8

Mann-Whitney U-Test Results on DV1 Test Scores

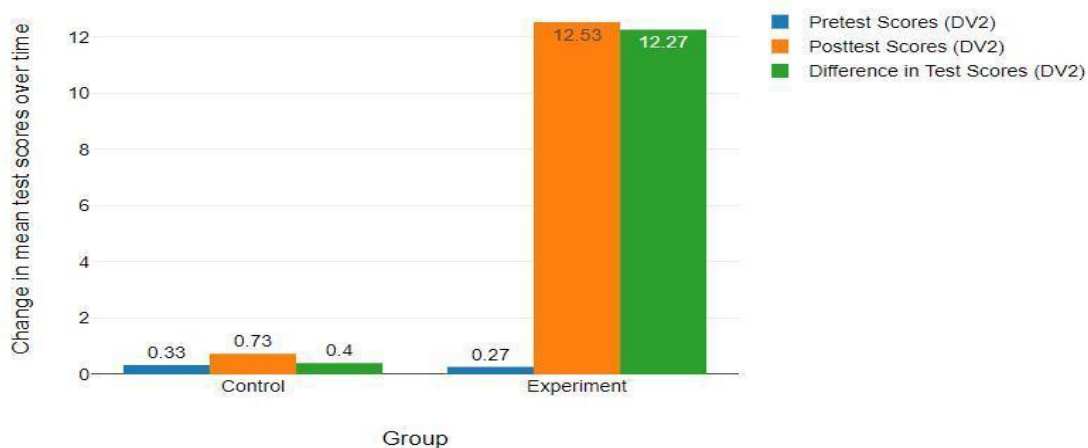
Variable	Control		Experiment		r	z	p
	n	Mean rank	n	Mean rank			
(DV1)	15	8.0	15	23.0	0.86	-4.73	<.001

(vi) Effect on Initial Letter Sound Proficiency (DV2)

Based on the pretests scores, post-test scores and their differences for the dependent variable, initial letter sound proficiency (DV2) significantly higher for the experimental group than the control group, as graphically indicated in Figure 5.

Figure 5

Group Comparison of Mean Test Scores for Initial Letter Sound Proficiency



A Shapiro-Wilk test was then performed and showed that the distribution of test scores for initial letter sound proficiency departed significantly from normality, $W(30) = .78$, $p < .001$. A Mann-Whitney U test was conducted to compare initial letter sound proficiency test scores between the experimental and control groups (see Table 9).

The results indicated that initial letter sound proficiency (DV2) test scores for the experimental group ($Mdn = 12.00$, mean rank = 23.00) were 99.97% significantly higher than those of the control group ($Mdn = 0.03$, mean rank = 8.00), $z = 4.76$, $p < .001$, with a large effect size, $r = 0.87$.

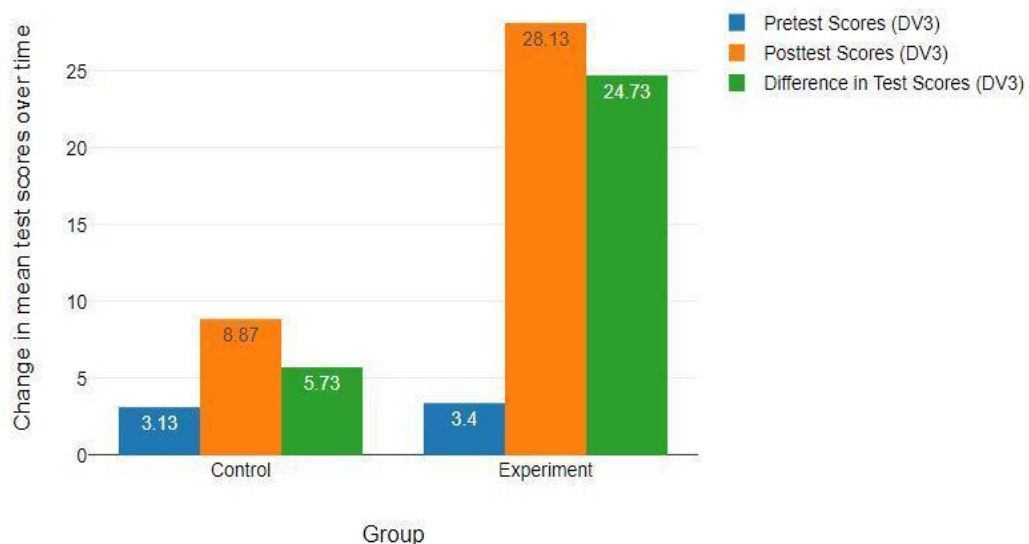
Table 9
Mann-Whitney U-Test Results on DV2 Test Scores

Variable	Control		Experiment		<i>r</i>	<i>z</i>	<i>p</i>
	<i>n</i>	Mean rank	<i>n</i>	Mean rank			
(DV2)	15	8.0	15	23.0	0.87	-4.76	<.001

(vii) Effect on vocabulary acquisition (DV3)

Based on the pretests scores, post-test scores and their differences for the dependent variable, vocabulary acquisition (DV3) significantly higher for the experimental group than the control group, as graphically indicated in Figure 6.

Figure 6
Group Comparison of Mean Test Scores for Vocabulary Acquisition



The pretest scores for the control group ($M = 3.13$, $SD = 1.46$, 95% $CI = 2.33, 3.94$) and the experimental group ($M = 3.4$, $SD = 1.24$, 95% $CI = 2.71, 4.09$) indicated that at the beginning of the study, the two groups were comparable in terms of their vocabulary acquisition abilities. Post-test scores indicated a significant increase in the experimental group compared to considerably lower post-test scores for the control group. The difference in test scores further highlighted the improvement in the experimental group ($M = 24.73$, $SD = 2.19$, 95% $CI = 23.52, 25.94$) over the control group ($M = 5.73$, $SD = 1.83$, 95% $CI = 4.72, 6.75$), indicating that the multisensory phonics curriculum had a positive impact on vocabulary acquisition (DV3).

A Shapiro-Wilk test was then performed and showed that the distribution of test scores for vocabulary acquisition departed significantly from normality, $W(30) = .78$, $p < .001$. A Mann-Whitney U test was conducted to compare vocabulary acquisition test scores between the experimental and control groups (see Table 10). The results indicated vocabulary acquisition (DV3) test scores for the experimental group ($Mdn = 25.00$, mean rank = 23.00) were 80% higher than those of the control group ($Mdn = 5.00$, mean rank = 8.00), $z = 4.70$, $p < .001$, with a large effect size, $r = 0.86$.

Table 10

Mann-Whitney U-Test Results on DV3 Test Scores

Variable	Control		Experiment		r	z	p
	n	Mean rank	n	Mean rank			
(DV3)	15	8.0	15	23.0	0.86	-4.70	<.001

Summary of Main Findings

This study investigated the effectiveness of a multisensory phonics curriculum in developing phonemic awareness among kindergarten-aged Chinese students learning English as a foreign language (EFL). The researchers employed a mixed-methods approach, combining quantitative and qualitative data to evaluate the program's impact. The main findings of this study can be summarised as follows:

(i) Effectiveness of the Multisensory Phonics Curriculum

Test Scores and Parental Feedback: The study found a strong positive correlation between student test scores and parental feedback. This suggests that improvements in phonemic awareness, as measured by the tests, aligned with parents' perceptions of their children's progress. Parental feedback also highlighted a preference for interactive activities, home resources, and progress updates.

Observational Findings: Observations in classrooms revealed a positive correlation between student engagement and participation in the learning process. Students in the experimental group (using the multisensory phonics curriculum) demonstrated higher levels of engagement, active participation, and interaction with learning materials compared to the control group. These observations support the theory that multisensory approaches can enhance student engagement and learning.

Test Score Correlations: The results showed a weak correlation between phonemic awareness skills in the pre-test scores, but a moderate to strong correlation in the post-test scores. This suggests that as students progressed through the phonics program, their phonemic awareness skills became more interconnected.

(ii) Impact on Specific Skills

Letter Recognition and Letter-Sound Recognition: The study found significant improvement in both letter recognition and letter-sound recognition among students in the experimental group compared to the control group. This indicates that the phonics curriculum was effective in teaching students to identify letters and their corresponding sounds.

Initial Letter Sound Proficiency: The experimental group also demonstrated a substantial increase in their ability to sound out initial letter sounds compared to the control group. This supports the effectiveness of the structured multisensory approach in developing this specific phonemic awareness skill.

Vocabulary Acquisition: The targeted multisensory phonics curriculum significantly enhanced vocabulary acquisition in the experimental group. This suggests that the program's focus on phonemic awareness and letter-sound recognition contributed to vocabulary growth.

Limitations of the Study

The study acknowledges that the findings, while promising, have some caveats. Here's a breakdown of the key points:

- **Limited generalizability:** The study involved a small group of Chinese EFL learners. The results might not apply to all EFL learners or those from different backgrounds. Cultural factors and prior exposure to English could influence outcomes. Future research needs larger, more diverse samples across different contexts.
- **Short study duration:** The 12-week program showed positive effects, but long-term impacts on phonemic awareness and vocabulary retention are unclear. Future studies with extended durations are needed.
- **Unobserved factors:** The study may not have accounted for all influences. Student variations in knowledge, learning styles, teacher effectiveness, and classroom dynamics could have played a role. Parental feedback, while valuable, is subjective and might be biased. Future research needs to consider these factors and use objective measures alongside parental reports.

- Technology usage: The positive aspects of technology in the experimental group (engagement, multisensory learning) need to be balanced with potential downsides of excessive screen time for young learners.
- Control group considerations: The control group received a language-integrated approach, which might be less effective in developing phonemic awareness compared to the multisensory phonics curriculum. Future studies could involve multiple control groups using different methods for a more rigorous comparison.
- Hawthorne effect: Increased attention and novelty of the intervention in the experimental group might have contributed to some of the observed improvements, not just the curriculum itself.

Conclusion

This study provided strong evidence that a multisensory phonics curriculum can be effective in developing phonemic awareness among kindergarten-aged Chinese EFL students. The findings offered valuable insights for educators, policymakers, and researchers interested in improving early language learning experiences.

Apart from that, the study's findings support educational theories like behaviourism, constructivism, and socio-cultural theory. The effectiveness of the multisensory approach aligns with the idea that students learn best by actively constructing their understanding through engaging activities. The mixed-methods approach, combining quantitative and qualitative data, provided a comprehensive evaluation of the program. This approach can serve as a model for future research in education.

The study suggests that educators should consider incorporating multisensory phonics approaches in their teaching. Additionally, involving parents, providing professional development for teachers, and offering tiered support for struggling learners are crucial aspects for successful implementation. Future research should explore the generalizability of the findings by including a more diverse sample and conducting comparative studies across different contexts. Further investigation is needed to identify the most effective elements of the multisensory approach and compare different phonics programs.

REFERENCES

- Agegnehu, A. Z., Bachore, M. M., & Ayele, Z. A. (2023). Effects of an explicit rime-based phonics method on the phonological awareness outcomes of Ethiopian EFL children. *Cogent Education*, 10(1), 2203582. <https://doi.org/10.1080/2331186X.2023.2203582>
- Attia, A. E. (2020). The effect of a multisensory synthetic phonics programme in developing automatic word recognition and spelling among young learners. *قيوبرتلا*, 28(1), 73-117 مولعلا.
- Bdeir, M., Bahous, R., & Nabhani, M. (2022). Improving reading readiness in kindergarten children through early phonological awareness interventions. *International Journal of Primary, Elementary and Early Years Education*, 50(3), 348-360. <https://doi.org/10.1080/03004279.2020.1851740>
- British Educational Research Association (BERA). (2018, June 20). *Ethical guidelines for educational research* (4th Ed.). <https://www.bera.ac.uk/publication/ethical-guidelines-foreducational-research-2018>
- Chen, M., Shing, S. R., Lun, W. W., Yusoff, N., Ping, C. K., & Shak, M. S. Y. (2022). Empirical studies on phonics in mainland China: A systematic review. *Frontiers in Economics and Management*, 3(5), 506-516.
- Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge. <https://doi.org/10.4324/9781315456539>
- Cope, B., & Kalantzis, M. (2015). Interpreting evidence-of-learning: Educational research in the era of big data. *Open Review of Educational Research*, 2 (1), 218-239.
- Creswell, J. W., & Guetterman, T.C. (2018). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (6th ed.). Pearson.
- Huo, S., & Wang, S. (2017). The effectiveness of phonological-based instruction in English as a foreign language students at primary school level: A research synthesis. *Frontiers in Education*, 2(15), 1-13. <https://doi.org/10.3389/educ.2017.00015>
- Jiang, Y., Gai, X., Üstün-Yavuz, M. S., Zhang, M., & Thomson, J. M. (2023). A meta-analysis of training effects on English phonological awareness and reading in native Chinese speakers. *PsyCh Journal*, 12(5), 599-617. <https://doi.org/10.1002/pchj.675>
- Johnson, A. P. (2019). *Essential learning theories: Applications to authentic teaching situations*. Rowman & Littlefield Publishers.
- Liu, X. (2022). Study on the popularity of phonics among Chinese elementary students and its implications for EFL teachers. In *2nd International Conference on Education: Current Issues and Digital Technologies (ICECIDT 2022)* (pp. 707-714). Atlantis Press.

Lund, T. (2012). Combining qualitative and quantitative approaches: Some arguments for mixed methods research. *Scandinavian Journal of Educational Research*, 56(2), 155-165.

Mertler, C. A. (2021). *Introduction to educational research*. Sage Publications

Onwuegbuzie, A. J., & Leech, N. L. (2004). Enhancing the interpretation of “significant” findings: The role of mixed methods research. *The Qualitative Report*, 9(4), 770-792.

Patten, M. L. (2017). *Understanding research methods: An overview of the essentials* (10th ed.). Routledge. <https://doi.org/10.4324/9781315213033>

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213-237. <https://doi.org/10.1075/lab.1.3.01par>

Pritchard, A. (2017). *Ways of learning: Learning theories for the classroom* (4th ed.). Routledge.

Shenoy, S., Iyer, A., & Zahedi, S. (2022). Phonics-based instruction and improvement in foundational reading skills of kindergartners in the Indian schooling context. *Early Childhood Education Journal*, 1-13. <https://doi.org/10.1007/s10643-022-01396-1>

Shin, J. K., & Crandall, J. J. (2018). *Teaching reading and writing to young learners*. In *The Routledge handbook of teaching English to young learners* (pp. 188-202). Routledge.

Sun, H., Steinkrauss, R., Van Der Steen, S., Cox, R., & De Bot, K. (2016). Foreign language learning as a complex dynamic process: A microgenetic case study of a Chinese child's English learning trajectory. *Learning and Individual Differences*, 49, 287-296.

Yeung, S. S., Siegel, L. S., & Chan, C. K. (2013). Phonological awareness, oral language proficiency and beginning reading development among Hong Kong Chinese kindergartners: An intervention study. *Reading and Writing*, 26(5), 681–704.